

ABSTRACT OF THE DISCLOSURE

A system and method for generating a synthetic text-to-speech TTS voice are disclosed. A user is presented with at least one TTS voice and at least one voice characteristic. A new synthetic TTS voice is generated by blending a plurality of existing TTS voices according to the selected voice characteristics. The blending of voices involves interpolating segmented parameters of each TTS voice. Segmented parameters may be, for example, prosodic characteristics of the speech such as pitch, volume, phone durations, accents, stress, mis-pronunciations and emotion.